We thank Reviewer #2 for the constructive review.

In view of the recommendation of two reviewers to restructure the entire paper, we only address the major comments. In the following, original comments of reviewer #2 are repeated in *italics.*

The paper focuses on the Benguela upwelling system using a global model and new biogeochemical parameterization scheme forced by 110 years of reanalysis forcing to diagnose deoxygenation and carbon trends in the region. The authors fail to qualify why they focus on the Benguela region. This seems like an odd choice for the questions posed,….

One goal of our study is to show that the representation of meso-scale structures does not only improve the regional biogeochemical mean state, but also affects the simulated long-term trends in an eastern upwelling system. A special focus is on the temporal evolution of the regional oxygen minimum zone. Such a centennial simulation with a global eddy-resolving model is not yet feasible. Therefore, we apply a stretched grid configuration to capture at least the meso-scale activity in the region of interest. Here, the Benguela region is an ideal candidate. By locating the poles on South America and South Africa, we get a high horizontal resolution with a feasible number of grid points due to the relative narrow South Atlantic.

…. *and they tread the NBUS and SBUS the same, but they are not the same. The authors need to cite more SBUS papers here – not a lot of representation of the SBUS system on the author team either.*

We agree that a more solid evaluation of the SBUS would have been appropriate and it will be provided in a new version of the manuscript. However, we do not understand the statement that we "*tread the NBUS and SBUS the same, but they are not the same*". From a modelling point of view, all processes represented in a biogeochemical model should be valid for any ocean region. None of our parameter settings are dependent on a geographical location. If a process is missing, it should become visible in one of the physical or biogeochemical tracer fields.

Their physical model is not adequate for the research questions. Physical model – agulhas leakage is weak as simulated – something that impacts one of their key points. The Indian Ocean in particular performs poorly in their streamfunction plot on Figure 4. The Agulhas Current is much too weak and too broad in their high resolution model (in fact, it looks better in the low resolution model).

We do not share this opinion of the reviewer at all. We agree that the Agulhas current is rather weak and broad, and that the retroflection is slightly shifted to the south in the high resolution set-up (higBUS). In contrast to the coarse resolution model, however, higBUS produces an Agulhas retroflection where a large bulk of water returns into the Indian Ocean and only a small proportion enters the South Atlantic.

As our focus is on long-term trends in the biogeochemistry, we need a model set-up which is applicable for multi-centennial simulations to achieve a consistent quasi-steady state initial condition. This comes with compromises between model resolution, performance, computer time demands, and available atmospheric forcing fields. E.g. regional models often apply the wind stress correction based on QuikSCAT data which are only available after 1999. Furthermore, transient simulations with two-way nested models are often initialized with spin-up runs of a couple of decades only (all spin-up integrations of the INALT family are 30 years long; Schwarzkopf etal, 2019, https://doi.org/10.5194/gmd-12-3329-2019). This spin-up time is much too short to reduce the model drift at depth and model results will be still close to the initial conditions for most tracers.

When running a model like this, that includes the Agulhas and the Agulhas retroflection, they need to be explicitly evaluated - the latter by looking at MADT for satellite and model - could be a mean state, and the former by using both satellite and in situ data - the ASCA and ACT data is available on the RSMAS website. While Indian Ocean features, they are highly relevant to the Benguela, especially the southern part, given the importance on water masses, transport of the Benguela, the generation of the shelf-edge jet, turbulence fields in the Cape Basin etc .. Also, there has been no effort to match their transport of the Agulhas against even published literature (e.g. Beal et al 2015: Capturing the transport variability of a western boundary Jet: results from the Agulhas Current Time-Series Experiment (ACT)). The Agulhas should be 70-80Sv, the 'mean' Benguela of the order of 20Sv.

We agree that a more solid evaluation of the Agulhas current system would have been beneficial. We will include your suggestions in the updated manuscript.

Their use of a climatological mean in Figure 8, used for an assessment of the coastal upwelling cells, doesn't make sense given that there is a strong seasonal signal in the south and a weak one further north. The southern Benguela upwelling cells will then of course be underestimated here. Given that their focus was on the coastal upwelling here, it is unclear why the authors took vertical velocities at a depth of 100 m when something more like 30/50 m would be more appropriate. Maybe the vertical layers in their model does not allow for this?

As indicated in the manuscript, the vertical grid resolution is 10-12m in the upper 100m. We compared monthly upwelling velocities at 50m and 100m depths and found basically the same variability (seasonal and interannual), although absolute velocities are of course slightly lower at 100m depth. From the biogeochemical point of view, 100m is typically considered the depth of the euphotic layer.
Furthermore, Figure 8 show a climatological mean of the months September to December. The model data presented thus capture the peak upwelling season in the southern BUS well, while slightly underestimating the northern part.

The description of Figure 7 made out that it captured the velocities well, however, their HR model shows intense offshore flow at ~34S which is just not realistic, and their offshore meridional velocities are underestimated, while they are too strong and uniform over the shelf. Instead of separating meridional and zonal velocities, one possible way forward is to show (as a proxy for surface geostrophic flow): either MADT (from satellite) compared with mean SSH from the models and/or current speeds overlaid with vectors.
Another metric for evaluation used in this region (given the intense offshore turbulence) is surface eddy kinetic energy (model vs. satellite). This would've given them a good sense of how well the leakage is reproduced.

We will carry out these analyses according to your suggestions.